PCP
PLANT & CELL PHYSIOLOGY

Explainable Deep Learning Reproduces a 'Professional Eye' on the Diagnosis of Internal Disorders in Persimmon Fruit

Takashi Akagi [1,†,*], Masanori Onishi[2,†], Kanae Masuda[1], Ryohei Kuroki[3], Kohei Baba[3], Kouki Takeshita[3], Tetsuya Suzuki[4], Takeshi Niikawa[4], Seiichi Uchida[3] and Takeshi Ise[5]

[1]Graduate School of Environmental and Life Science, Okayama University, Okayama, 700-8530 Japan
[2]Graduate School of Agriculture, Kyoto University, Kyoto, 606-8502 Japan
[3]Kyusyu University, Fukuoka, 819-0395 Japan
[4]Gifu Prefectural Agricultural Technology Center, Gifu, Japan
[5]Field Science Education and Research Center, Kyoto University, Kyoto, 606-8502 Japan
[†]These authors contributed equally to this work.
*Corresponding author: E-mail, takashia@okayama-u.ac.jp; Fax, +81-86-251-8337.
(Received 17 February 2020; Accepted 18 August 2020)

Recent rapid progress in deep neural network techniques has allowed recognition and classification of various objects, often exceeding the performance of the human eye. In plant biology and crop sciences, some deep neural network frameworks have been applied mainly for effective and rapid phenotyping. In this study, beyond simple optimizations of phenotyping, we propose an application of deep neural networks to make an image-based internal disorder diagnosis that is hard even for experts, and to visualize the reasons behind each diagnosis to provide biological interpretations. Here, we exemplified classification of calyx-end cracking in persimmon fruit by using five convolutional neural network models with various layer structures and examined potential analytical options involved in the diagnostic qualities. With 3,173 visible RGB images from the fruit apex side, the neural networks successfully made the binary classification of each degree of disorder, with up to 90% accuracy. Furthermore, feature visualizations, such as Grad-CAM and LRP, visualize the regions of the image that contribute to the diagnosis. They suggest that specific patterns of color unevenness, such as in the fruit peripheral area, can be indexes of calyx-end cracking. These results not only provided novel insights into indexes of fruit internal disorders but also proposed the potential applicability of deep neural networks in plant biology.

Keywords: Artificial intelligence • Backpropagation • Convolutional neural network • Image diagnosis • Physiological disorder.

Introduction

Noninvasive prediction of internal disorders in plants is an important issue for plant phenotyping, in both the experimental and agricultural fields. Attempts have been made to capture the indexes of internal reaction/responses from many viewpoints, where recent interest might be moving toward molecular regulation, including RNA-based characterization potentially utilized as biomarkers. Notwithstanding, 'empirical' prediction of the symptoms, with expertise and professional skills, can still produce better diagnostic results in most cases. The problem is more serious in agricultural sciences, where noninvasive and rapid prediction of internal disorders with reasonable costs would be indispensable for actual application. Such 'empirical' knowledge would, however, be attained only after several decades of technical experience, and it would be difficult to explain simply the regions of the image which are relevant to the diagnosis.

Object recognition and classification by deep neural networks have made rapid and widespread progress in the past few years. The Large Scale Visual Recognition Challenge (ILSVRC) (Russakovsky et al. 2015) based on the ImageNet dataset (Deng et al. 2009) has promoted development of various deep convolutional neural network (CNN) models involving visualization-related issues (Simonyan et al. 2013, He et al. 2015, Szegedy et al. 2015, Singh et al. 2016, Singh et al. 2018). Importantly, in 2015, a deep neural network model, named ResNet50, was able to exceed the human standard in image classification (He et al. 2015). This situation might imply that deep neural networks can promptly reproduce professional eyes to recognize specific objects, by learning the images 'empirically'. For plants, deep neural network models have been successfully applied mainly to detect symptoms of stresses/diseases in many plant species (Ramcharan et al. 2017, Ferentinos 2018, Ghosal et al. 2018, Singh et al. 2018). On the other hand, only a few applications of deep neural networks to predict internal (or invisible) characters in plants organs have been reported, such as for detection of internal physical damage of blueberry fruits (Wang et al. 2018). For object classification in plants, deep neural networks have been applied for the identification of the species depending on subtle differences, such as moss species (Ise et al. 2017). A big issue in object recognition or classification by deep neural networks was that we could not detect or characterize the regions of the image that are highly relevant to derive the result. To solve

Plant Cell Physiol. 61(11): 1967–1973 (2020) doi:10.1093/pcp/pcaa111, Advance Access publication on 26 August 2020, available online at https://academic.oup.com/pcp

this issue, various feature visualization methods, such as Layer-wise relevance propagation (LRP) (Bach et al. 2015) and its derivates (Iwana et al. 2019), Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al. 2016) and Guided backpropagation (Springenberg et al. 2015), have been proposed recently. Application of these methods to physiological analyses would allow cell- or site-specific analyses in early reaction steps.

Here, we exemplified prediction of an internal disorder in persimmon fruits. Persimmon is one of the major fruit crops, especially in East Asia. Similar to other fruit crops, persimmon fruit has some internal disorders, represented by quick softening or internal cracking, which substantially reduce the quality of fruit. Calyx-end cracking, also called 'hetasuki' in Japanese, is a typical physiological disorder of persimmon (**Fig. 1**). This internal disorder has been well classified into five grades (Yamada et al. 1988), in which higher degrees not only visually spoil the fruit quality, but often become a trigger of quick softening (Yamada et al. 1988, Yamada et al. 2002). There are some genetic or environmental factors affecting the incidence of the calyx-end cracking, whereas for each fruit, clear outer symptoms or biomarkers have not yet been identified. Importantly, it is suggested that only a limited number of experts in the field of fruit selection, who have been engaged in the actual cultivation of persimmon for several decades, can distinguish calyx-end cracking from its external appearance (**Supplementary Fig. S1**). Here, with application to this internal disorder in persimmon, we examined the utility of deep neural networks and their backpropagation for diagnosis of internal reactions in fruit.

Results

Deep neural networks allowed binary classification of calyx-end cracking

The experimental flow of this study is summarized in **Fig. 1**. A total of 3,173 persimmon fruits from cv. Fuyu were harvested at the fully ripened stage, in late November 2017, at Gifu city (N35.441721, E136.699894), Japan. Images were taken from the fruit apex side using a digital camera (NIKON COOLPIX P520), with a uniform gray background (**Supplementary Fig. S2**; see the Materials and Methods for details of the conditions). The degree of calyx-end cracking in persimmon fruits is categorized into five grades (level 0–4) (Yamada et al. 1988, Yamada et al. 2002; **Supplementary Table S1**). Two experts on the selection of persimmon fruits classified the grade of calyx-end cracking by visual observation of dissected fruit according to Yamada et al. (1988). Here, we attempted to use deep neural networks for classifying positive–negative (binary) categories using various thresholds for the cracking levels, employing

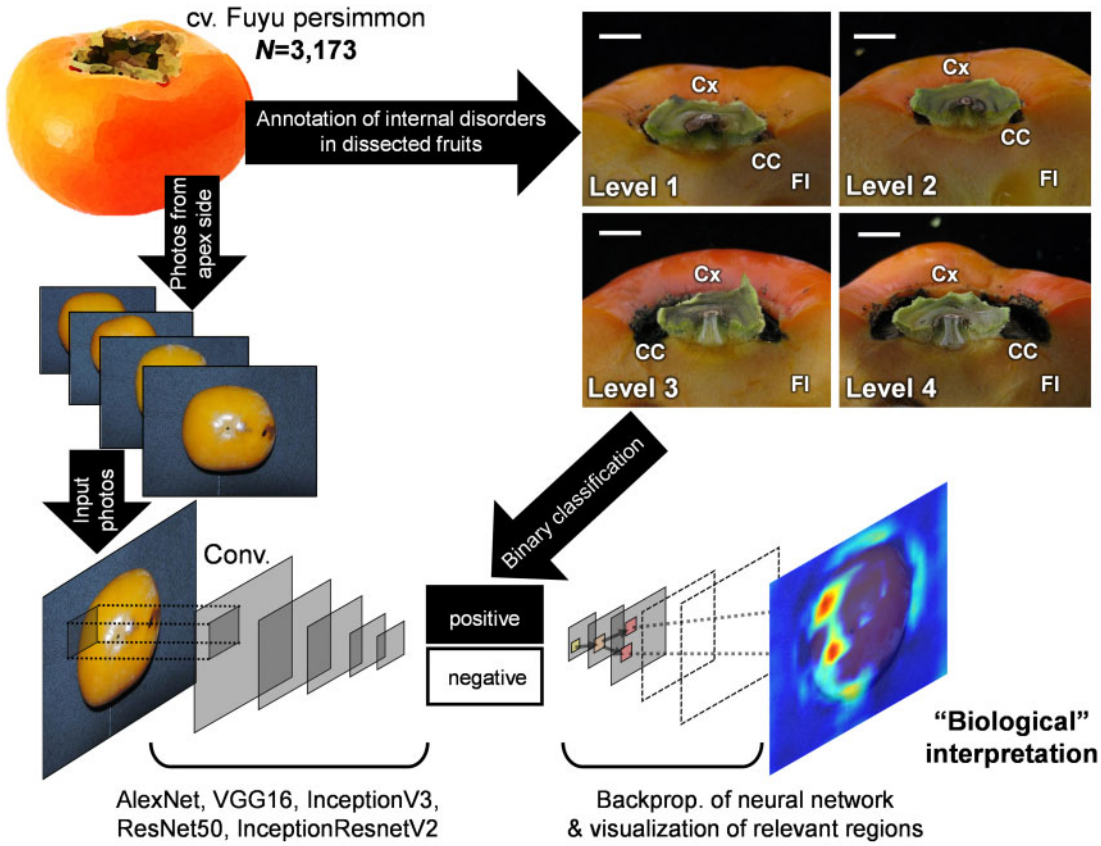

Fig. 1 Flow of the diagnosis and its visual explanation of calyx-end cracking in persimmon. Images from the apex side of a total of 3,173 cv. Fuyu persimmon fruits were subjected to CNN deep learning analysis to classify them into two categories of various thresholds of the degree of cracking. Backpropagation of the well-trained models allowed a visual explanation of the diagnosis, which renders 'biological' interpretation of the symptom of calyx-end cracking. CC, calyx-end cracking; Cx, dissected calyx; Fl, fruit flesh.

AlexNet, VGG16, InceptionV3, ResNet50 and InceptionResnetV2, implemented in Keras 2.2.4 (**https://keras.io/**) and pre-trained with ImageNet (**http://www.image-net.org/**) (see **Supplementary Table S2** for the detailed settings).

All of these five neural networks make good classifications of calyx-end cracking and the control in most positive/negative categorizations except level 0–3 (as negative) vs. level 4 (as positive) (**Table 1**, **Supplementary Fig. S3**). This exception in level 0–3 vs. level 4 would be mainly due to the imbalance being too severe, because we have only 65 fruits with level 4 calyx-end cracking, as discussed later. Among the neural network models, InceptionResnetV2 gave the highest accuracy (<90%) in the classification of the calyx-end cracking and the control, although even AlexNet, with only eight simple layers, showed reasonable classification with >75% accuracy for any categorizations (**Table 1**). With InceptionResnetV2, evaluations of classification performance with confusion matrix, distribution of predictions, receiver operating characteristic (ROC) curve and precision recall (PR) curve (**Fig. 2**) supported that the model could make good predictions in any categorizations [area under curve (AUC) value >0.77], while classification of level 0–1 vs. level 2–4 or level 0 vs. level 1–4 would be reliable for actual social/experimental implementation, due to the shape of the ROC and the prediction distribution.

In the models classifying level 0 and level 1–4, and level 0–1 and level 2–4, the predictions tended to be increased along with the levels of calyx-end cracking (**Fig. 3**). Levels 1 and 2 in particular showed statistically significant differences in the distribution of the predictions in both models ($P = 0.37e^{-8}$ and 0.0022 for models classifying 0 vs. 1 and 1 vs. 2, respectively). This situation is

Table 1 Prediction accuracy in training samples, in each model with four binary classification tasks

Model	Binary classification of calyx-end cracking levels			
	0 vs. 1–4	0–1 vs. 2–4	0–2 vs. 3–4	0–3 vs. 4
AlexNet	0.75	0.87	0.85	0.99a
VGG16	0.79	0.89	0.83	0.99a
InceptionV3	0.81	0.87	0.89	0.99a
ResNet50	0.81	0.91	0.85	0.99a
InceptionResnetV2	0.79	0.89	0.86	0.99a

[a]Due to the severe imbalance between level 0–3 and level 4. Level 0–3 occupied >98% of the samples.

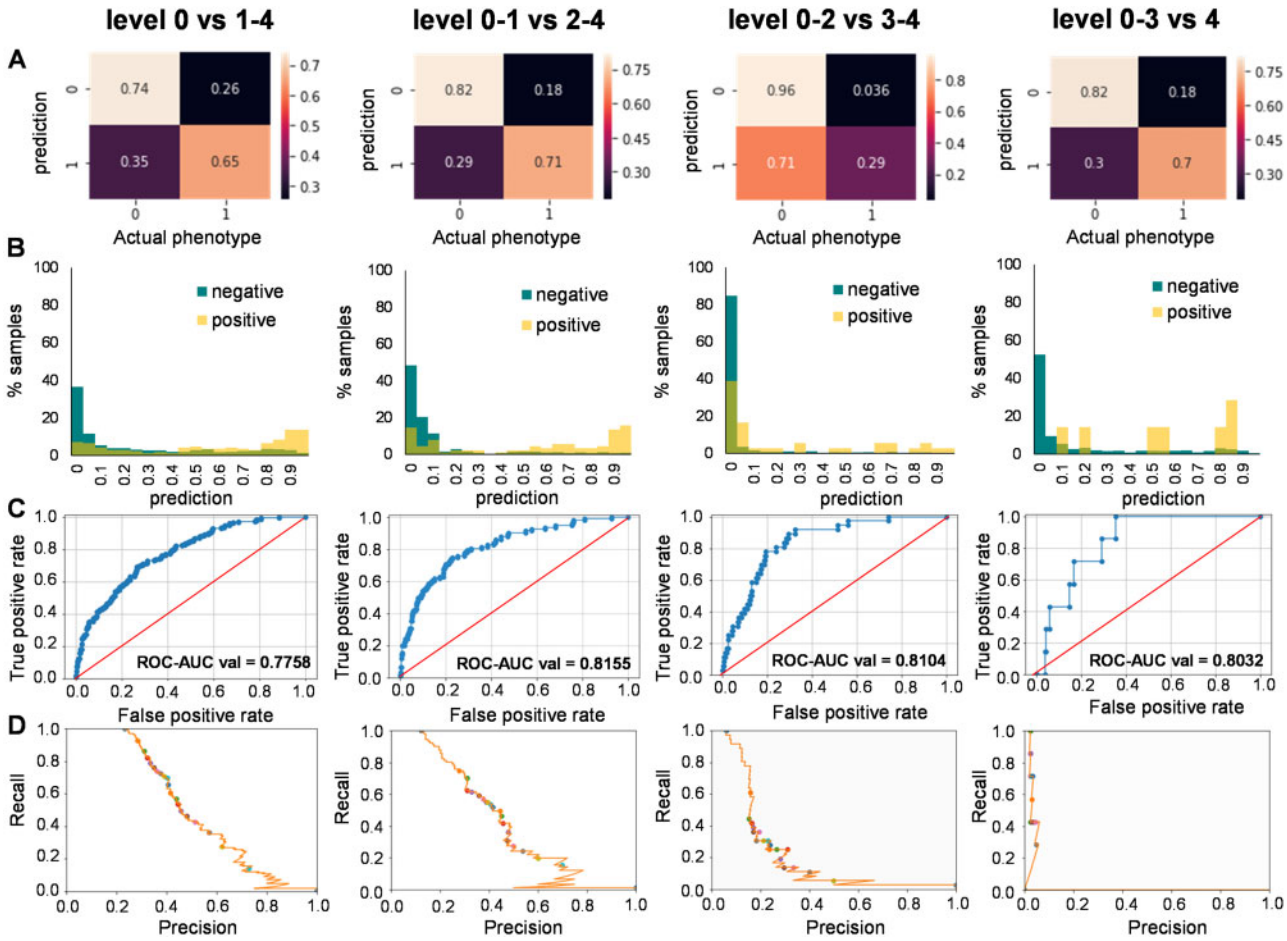

Fig. 2 Evaluation of the trained models in each categorization. Confusion matrix (A), distribution of prediction (B), ROC curve (C) and PR curve (D) in the four categorizations of binary classification, level 0 vs. 1–4, level 0–1 vs. 2–4, level 0–2 vs. 3–4 and level 0–3 vs. 4. Although no substantial differences were observed in ROC-AUC values and in confusion matrix among the categorizations, the prediction distribution and PR curves suggested that categorization of level 0–1 vs. 2–4 could construct the most reliable model.

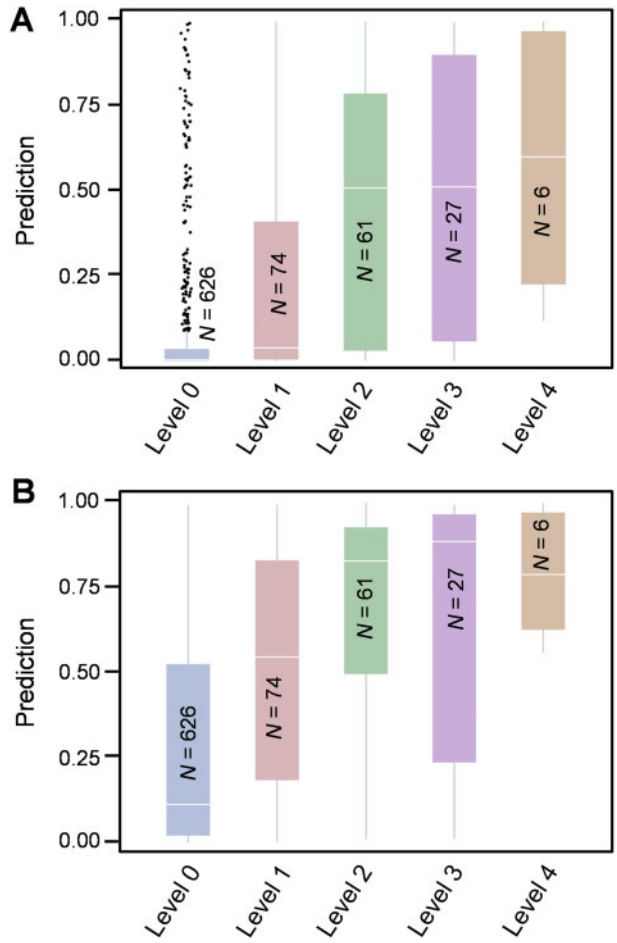

Fig. 3 Comparison of the prediction in the five degrees of calyx-end cracking. The models trained with the categorizations of level 0 vs. level 1–4 (A) and level 0–1 vs. level 2–4 (B) were used for the detection of prediction in each level. There is a clear tendency that prediction was increased according to the cracking level. In both models, statistically significant differences ($P < 0.01$) were detected between level 0 and level 1 (and >1) and between level 1 and level 2 (and >2), by Student's *t*-test.

reminiscent of that which we would be able to apply in regression on this internal disorder.

Visualization of 'the reason for diagnosis'

To visualize the image regions relevant to the diagnosis of calyx-end cracking, we applied the following feature visualization methods: Guided backpropagation (Springenberg et al. 2015), two LRP derivates [LRP-Epsilon, and LRP-Sequential B which uses LRP-Alpha/Beta for convolution layers and LRP-Epsilon for fully connected layers (Iwana et al. 2019)], Grad-CAM (Selvaraju et al. 2016, Selvaraju et al. 2017) and Guided Grad-CAM (Selvaraju et al. 2017). The neural network models with highly complicated layer structures, such as InceptionsV3, ResNet50 or InceptionResnetV2, which tended to show higher classification performance (**Table 1**), are thought to be harder to apply to backpropagations to the upper layers. Here, to characterize the performance and the potential applicability of backpropagations, we mainly selected a simple model, VGG16, which showed sufficient performance in classification of calyx-end cracking (**Table 1**).

We visualized the regions of the images relevant to the diagnosis of 93 test images by each feature visualization method. The highly relevant regions by Grad-CAM were too coarse to specify the relevant regions at the pixel level (**Fig. 4A, B** for four images of each 0–3 cracking level; **Supplementary Fig. S4** for a further 10 images). On the other hand, Guided backpropagation, LRP-Sequential B, LRP-Epsilon and Guided Grad-CAM provided finer heatmaps (**Fig. 4C–F**, **Supplementary Fig. S4**). However, the relevant regions were often inconsistent among the feature visualization methods. To analyze the regions with higher relevance to the diagnosis of calyx-end cracking, the distribution of the relevance values along with the distance from the outer contour of the fruit was calculated by using all validation images. As shown in **Fig. 5A–D**, the distribution is shown as a two-dimensional histogram $H(r, d)$, where r is the relevance level of a pixel and d is the normalized distance of the pixel from the contour. Higher relevance levels tend to be located mainly around the apex (approximately $d \in [0.75,1.0]$) in all feature visualization methods. The higher relevant levels are also found on peripheral (i.e. near-contour) parts of fruit (approximately $d \in [0,0.2]$) (**Fig. 5A, B, D**), except in LRP-Epsilon (**Fig. 5C**). It is noteworthy that these objective analysis results were at least partially consistent with 'conventional empirical diagnosis', where color unevenness in the pericarp, especially in the area close to the calyx (or peripheral parts in pictures from the fruit apex side), is thought to be an index of certain stresses (Yamada et al. 1988). In reality, dissection of the fruit with severe cracking suggested that reddish coloration was extended from the calyx end to the peripheral area, on the cracking side (**Fig. 5E**). On the other hand, substantial relevance peaks around the apex (**Fig. 5A–D**) would not be interpretable from conventional empirical diagnosis; this result may provide novel insights into physiological reactions in the signaling of internal disorders.

Discussion

In this study, we applied only approximately 3,000 images of the fruits to successfully predict an internal disorder and visualize the index of the disorder. This is only 'an example' to show the potential of deep neural networks to predict physiological reactions, potentially exceeding the performance of the eyes of the professional. In reality, calyx-end cracking is very hard to predict from the outer appearance of the fruit, even with the eyes of professionals honed over decades, while simple neural network models could classify them well and unveil 'the reason for the classification', as shown in this study. This would suggest that we might be able to access more detailed or earlier physiological reactions by combining deep neural networks and omics approaches. For instance, rapid prediction and visualization of the symptoms of disorders in plant organs or of regeneration in calli would enable transcriptomic or metabolomic approaches specific to the tissues with the symptoms, and characterization of the early reactions in these specific tissues. The

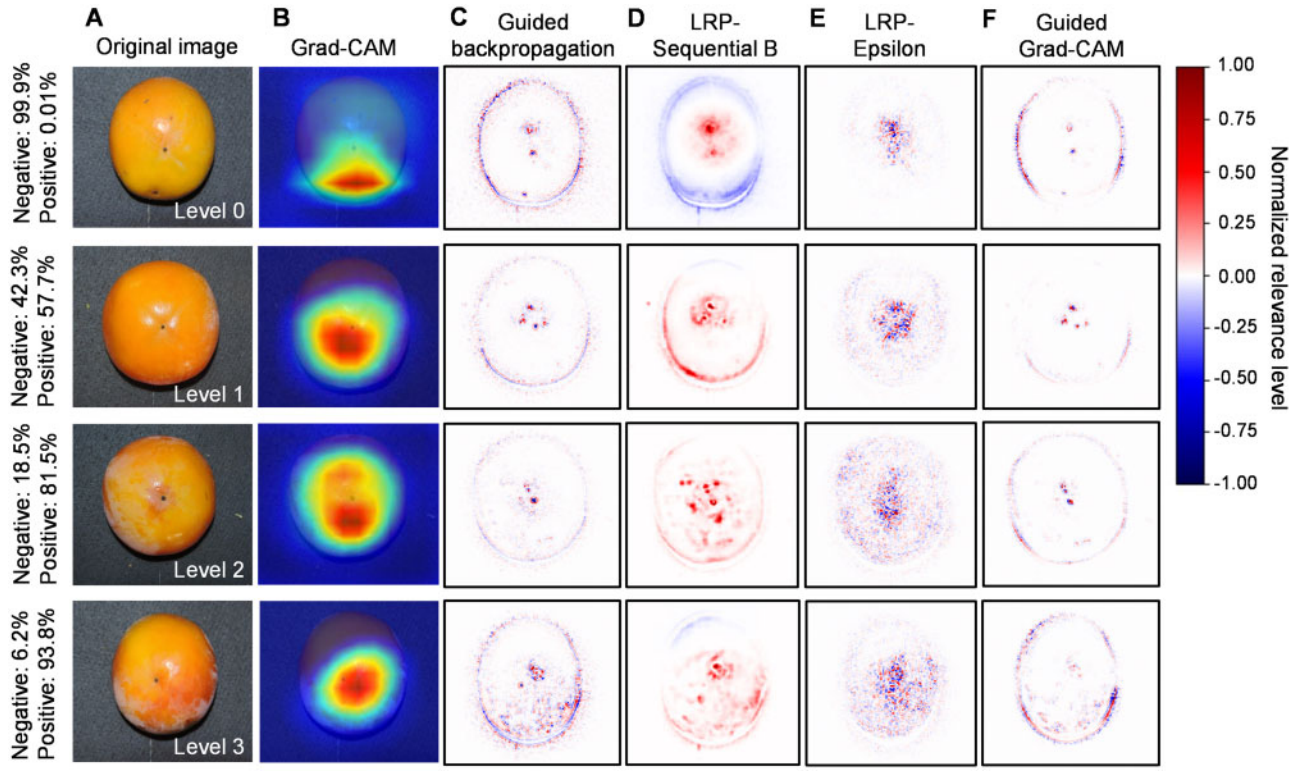

Fig. 4 Visualization of explanatory factors in the prediction of the regions relevant to the diagnosis of calyx-end cracking. (A) Original images of the fruits with level 0–3 calyx-end cracking. The relevant regions in the VGG16 model were visualized with Grad-CAM (B), Guided backpropagation (C), LRP-Sequential B (D), LRP-*E*psilon (E) and Guided Grad-CAM (F). For Grad-CAM (B), relevant regions in the block5_conv3 (immediately before the fully connected layer) are visualized very coarsely, while they still capture the areas with color unevenness in the samples with high levels of calyx-end cracking (levels 2 and 3). The other four feature visualization methods (C–F) exhibited finer relevant regions, although they were not completely consistent among the methods. In particular, in the samples with high cracking levels (levels 2 and 3), these methods commonly visualized relevant regions around the areas with color unevenness.

symptoms would not be limited to visual images, but any signals which can be converted into arrays, such as the spectrum of absorbance in chemicals, would be applicable to deep neural networks and their backpropagation.

For actual use of deep learning in biological experiments, imbalances among the classes would be one of the biggest issues in any models, since the 'case' (or presence) would normally be rarer than the controls. We implemented class weights (Sarafianos et al. 2018) to overcome the class imbalances of calyx-end cracking. In our classification of the presence/absence of cracking, at least 75% of the samples were labeled as 'no symptom' (absence). This situation would substantially reduce the prediction accuracy if we had no information regarding the class weights (**Supplementary Fig. S5**). The better prediction would also depend on the fitting of learning rates (Sutskever et al. 2013). In our analysis, change of the learning rate in some models could result in different classification performance, although the trends are dependent on the models (**Supplementary Table S3**).

In the comparison of the characteristics of backpropagations, biological experiments may often need to visualize finer relevant regions (or contributing parts) than those outputted from Grad-CAM, to narrow down the explanatory cells/organs, while Grad-CAM can visualize weights in each layer (Selvaraju et al. 2016). On the other hand, other backpropagations used in this study could visualize very fine relevant regions in the image in this study. Hence, in the application of feature visualization techniques to plant sciences, proper selection of methods depending on situations might be a key to allow 'novel biological interpretations', as used in molecular biological modeling (Hochuli et al. 2018) or in the characterization of stress-induced symptoms in plants (Fujita et al. 2016). A visual explanation from multiple aspects would shed light on the hidden explanations or indexes of internal disorders in plants, and propose novel interpretations and analytical approaches.

Materials and Methods

Assessment of calyx-end cracking in persimmons

Persimmon cv. Fuyu has been maintained in the Agricultural Technology Center, Gifu. A total of 3,173 fully matured fruits from four (38, 57, 58 and 92 years) trees were located on a gray background sheet to take photos from the fruit apex side, according to conventional fruit sorting for marketing, using a digital camera (NIKON COOLPIX P520) (**Supplementary Fig. 2**). The exposure and white balance were default (auto), and the size of the image was 1,600 × 1,200 pixels. The fruits were dissected to visually categorize the degrees of calyx-end cracking into five levels (**Supplementary Table S1**) by two experts in Gifu Prefectural Agricultural Technology Center, who have been trained for

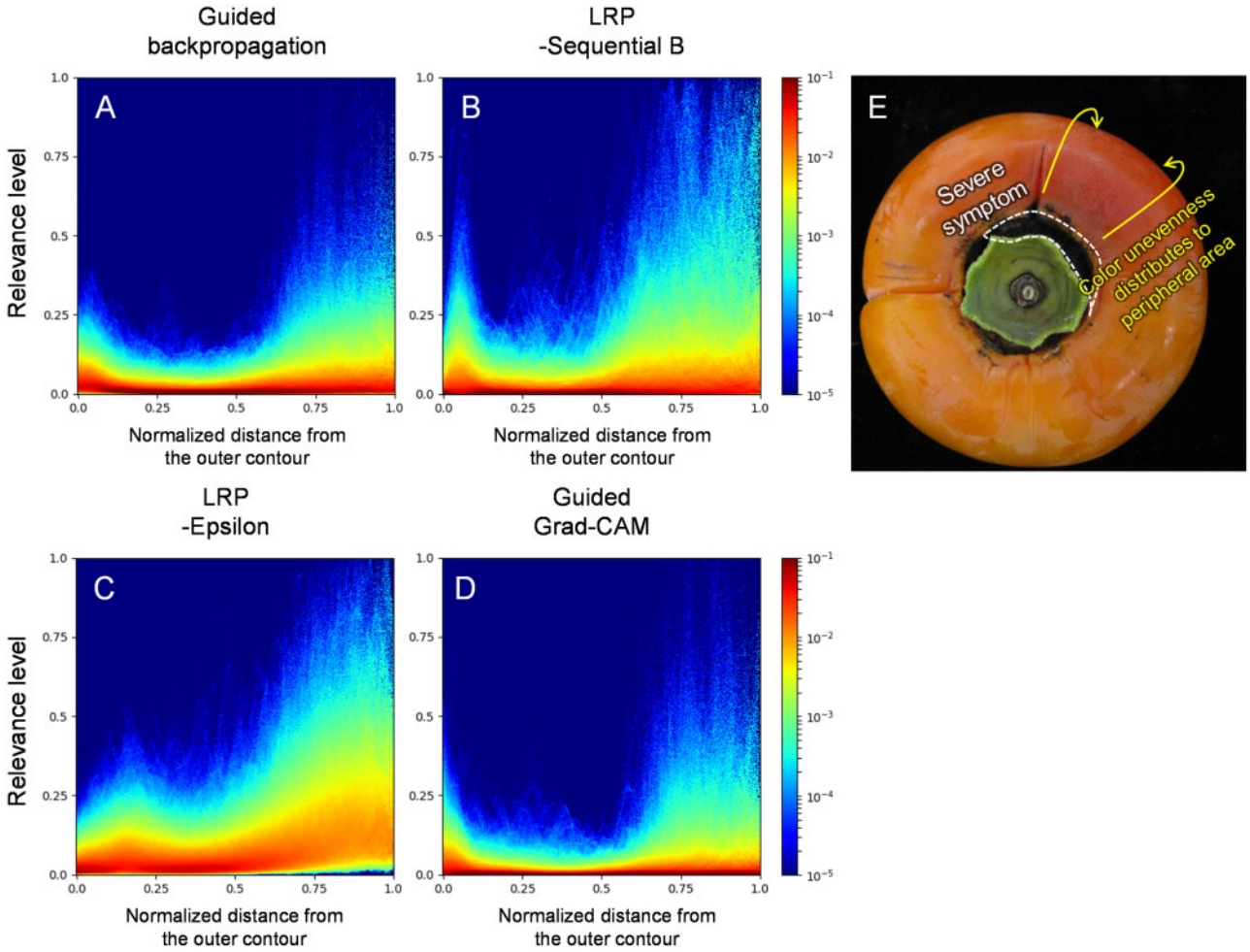

Fig. 5 Distribution of relevance levels in fruits. (A–D) Two-dimensional histogram $H(r, d)$, where r is the relevance level of a pixel and d is the normalized distance of the pixel from the contour, by four feature visualization methods: Guided backpropagation (A), LRP-Sequential B (B), LRP-*E*psilon (C) and Guided Grad-CAM (D). Higher relevance levels tend to be located mainly around the apex ($d \in [0.75,1.0]$) in all feature visualization methods. The higher relevant levels are also found on peripheral (i.e. near-contour) parts of the fruit (approximately $d \in [0,0.2]$) (A, B and D), except in LRP-Epsilon (C). Observation of a dissected fruit with partially severe calyx-end cracking, from the calyx side (E), might support this distribution. Only the side with severe cracking showed color unevenness expanding to the peripheral part, probably due to the stress responses and the signaling.

selection of cv. Fuyu persimmon fruit for >20 years. For access to the original picture data, contact the corresponding author.

Image processing and construction of the neural networks

We resized each image sample to the size of 224 × 224 pixels and then augmented them by flipping them horizontally and vertically and by rotation. The image data were rescaled in the range of 0–1. We randomly selected 70% of images for training and 30% for testing. We use the standard deep neural network models called VGG16, InceptionsV3, ResNet50 and InceptionResnetV2. These models are implemented in Keras 2.2.4 (**https://keras.io/**). AlexNet was manually constructed according to Krizhevsky et al. (2012) with the sequential model API of Keras. Each model was pre-trained with the standard image data set called ImageNet (**http://www.image-net.org/**), and then, their fully connected layer was customized for binary classification. We used the class weight option, which is prepared in Keras for balancing the two categories with different sample numbers. Training and testing with those models were run on Ubuntu 18.04 (DeepStation DK1000, 16GB RAM, GPU = 1). The detailed setting of the neural networks is summarized in **Supplementary Table S2**. The performance of the trained models was evaluated with the distribution of predictions, confusion matrix, ROC-AUC values and PR curves in the test samples. The differences in prediction distribution among the calyx-end levels were examined with Student's *t*-test.

Feature visualization methods

Feature visualization methods reveal the input image regions that are highly relevant to the final classification. An implementation of the following feature visualization methods using the iNNvestigate library (Alber et al. 2019) can be found at **https://github.com/uchidalab/softmaxgradient-lrp**; this implementation was used in our experiments. Guided backpropagation, Grad-CAM and Guided Grad-CAM are similar because they commonly assume that relevant regions give a larger gradient (i.e. a large impact) to the class likelihood by the deep neural networks. Guided backpropagation (Springenberg et al. 2015) tries to find the input pixels that have a large impact on the class likelihood using the orthodox backpropagation procedure (used for training the networks), except that the gradient value in the procedure undergoes the rectified-linear function to change the negative gradient values to zero. Grad-CAM (Selvaraju et al. 2017) tries to find the high-impact regions not in the input image but in the feature map by the last convolutional layer. Since the feature map is smaller than the input image, it is necessary to enlarge the relevance map to obtain the size of the original input image—this is the reason why it can only give coarse

visualizations. Guided Grad-CAM gives its visualization by simply multiplying Guided backpropagation and Grad-CAM to utilize the merits of both methods. Different from the above gradient-based visualization methods, LRP and its variants (Montavon et al. 2019) try to reveal the input pixels that are highly relevant to the results by decomposing the class likelihood into the input pixels. This decomposition is made by applying a simple weighted reallocation rule from the output layer to the input layer, where the weights are the same as those of the individual neurons. The variants, such as LRP-Sequential B and LRP-Epsilon, are derived by slightly modifying the parameters in the reallocation rule.

Distribution of relevance levels in fruit

The distribution, i.e. the two-dimensional histogram $H(r, d)$ (**Fig. 5A–D**), is derived by the following procedure (**https://github.com/Takeshiddd/persimmon_GuidedGradCAM_visutalize**). First, the fruit region is extracted from each fruit image using color clustering after increasing the color saturation. Since simple binarization fails to have the accurate fruit region around the shaded part of the fruit, color clustering was employed with a k-means algorithm ($k = 2$). Second, mathematical morphology operations are applied to the fruit region image to remove noisy regions. Third, distance transformation is applied to determine the distance from the fruit contour at each pixel. The distance is normalized so that the maximum distance becomes 1. Finally, the two-dimensional histogram $H(r, d)$ is obtained by counting the number of pixels with the relevance level r (given by a feature visualization method) and the normalized distance d.

Supplementary data

Supplementary data are available at PCP online.

Funding

PRESTO from the Japan Science and Technology Agency (JST) [JPMJPR15Q1] to T.A.; a Grant-in-Aid for Scientific Research on Innovative Areas from JSPS [19H04862] to T.A.; and a Grant-in-Aid for JSPS Fellows [19J23361] to K.M.

Disclosures

The authors have no conflicts of interest to declare.

References

Alber, M., Lapuschkin, S., Seegerer, P., Hägele, M., Schütt, K.T., Montavon, G., et al. (2019) INNvestigate neural networks! *J. Machine Learn. Res.* 20: 1–8.

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R. and Samek, W. (2015) On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS One* 10: e0130140.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L. (2009) Imagenet: a large scale hierarchical image database. *In 2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, pp. 248–255.

Ferentinos, K.P. (2018) Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* 145: 311–318.

Fujita, E., Kawasaki, Y., Uga, H., Kagiwada, S. and Iyatomi, H. (2016) Basic investigation on a robust and practical plant diagnostic system. *In* 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, pp. 989–992.

Ghosal, S., Blystone, D., Singh, A.K., Ganapathysubramanian, B., Singh, A. and Sarkar, S. (2018) An explainable deep machine vision framework for plant stress phenotyping. *Proc. Natl. Acad. Sci. USA* 115: 4613–4618.

He, K., Zhang, X., Ren, S. and Sun, J. (2015) Deep residual learning for image recognition. *In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, pp. 770–778.

Hochuli, J., Helbling, A., Skaist, T., Ragoza, M. and Koes, D.R. (2018) Visualizing convolutional neural network protein–ligand scoring. *J. Mol. Graph. Model.* 84: 96–108.

Ise, T., Minagawa, M. and Onishi, M. (2017) Identifying 3 moss species by deep learning, using the 'chopped picture' method. *arXiv* preprint 1708.01986.

Iwana, B.K., Kuroki, R. and Uchida, S. (2019) Explaining convolutional neural networks using softmax gradient layer-wise relevance propagation. *arXiv* preprint 1908.04351.

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) ImageNet classification with deep convolutional neural networks. *Adv. Neural Inform. Process. Sys.* **http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf**.

Montavon, G., Binder, A., Lapuschkin, S., Samek, W. and Müller, K.R. (2019) Layer-wise relevance propagation: an overview. *In Explainable AI: Interpreting, Explaining and Visualizing Deep Learning. Lecture Notes in Computer Science. Edited by Samek, W., Montavon, G., Vedaldi, A., Hansen, L.K., Müller, K.-R.*, Vol. 11700. pp. 193–209. Springer, Cham.

Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J. and Hughes, D.P. (2017) Deep learning for image-based cassava disease detection. *Front. Plant Sci.* 8: 1852.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015) ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* 115: 211–252.

Sarafianos, N., Xu, X. and Kakadiaris, I.A. (2018) Deep imbalanced attribute classification using visual attention aggregation. *In Computer Vision—ECCV 2018*. Edited by Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. pp. 680–697. Springer.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D. (2017) Grad-CAM: visual explanations from deep networks via gradient-based localization. *In* 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017. pp. 618–626.

Selvaraju, R.R., Das, A., Vedantam, R., Cogswell, M., Parikh, D. and Batra, D. (2016) Grad-CAM: why did you say that? arXiv preprint 1611.07450.

Simonyan, K., Vedaldi, A. and Zisserman, A. (2013) Deep inside convolutional networks: visualising image classification models and saliency maps. *arXiv* preprint 1312.6034.

Singh, A., Ganapathysubramanian, B., Singh, A.K. and Sarkar, S. (2016) Machine learning for high-throughput stress phenotyping in plants. *Trends Plant Sci.* 21: 110–124.

Singh, A.K., Ganapathysubramanian, B., Sarkar, S. and Singh, A. (2018) Deep learning for plant stress phenotyping: trends and future perspectives. *Trends Plant Sci.* 23: 883–898.

Springenberg, J.T., Dosovitskiy, A., Brox, T. and Riedmiller, M. (2015) Striving for simplicity: the all convolutional net. *arXiv* preprint 1412.6806.

Sutskever, I., Martens, J., Dahl, G. and Hinton, G. (2013) On the importance of initialization and momentum in deep learning. *In* Proceedings of the 30th International Conference on Machine Learning, Atlanta, Georgia, USA. pp. 1139–1147.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015) Going deeper with convolutions. *In* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2015. pp. 1–9.

Wang, Z., Hu, M. and Zhai, G. (2018) Application of deep learning architectures for accurate and rapid detection of internal mechanical damage of blueberry using hyperspectral transmittance data. *Sensors* 18: 1126.

Yamada, M., Ikeda, I., Yamane, H. and Hirabayashi, T. (1988) Inheritance of fruit cracking at the calyx end and stylar end in Japanese persimmon. *J. Japan. Soc. Hortic. Sci.* 57: 8–16.

Yamada, M., Sato, A. and Ukai, Y. (2002) Genetic differences and environmental variations in calyx-end fruit cracking among Japanese persimmon cultivars and selections. *Hortscience* 37: 164–167.